The Role of Spatial Mixing in the Spread of Foot-and-Mouth Disease

G. Chowell[1,2] *, A. L. Rivas[1], N. W. Hengartner[3],

J. M. Hyman[2], C. Castillo-Chavez[4]

[1] Department of Biological Statistics and Computational Biology, Cornell University,

432 Warren Hall, Ithaca, NY 14853, USA;

[2] Theoretical Division (MS B284), [3] Statistical Science (MS F600),

Los Alamos National Laboratory,

Los Alamos, NM 87545, USA;

[4] Department of Mathematics and Statistics, Arizona State University,

P.O. Box 871804, Tempe, AZ 85287-1804, USA

August 31, 2005

*Corresponding author. Email: gchowell@t7.lanl.gov Fax: (505) 665-5757

Abstract

A model of epidemic dispersal (based on the assumption that susceptible cattle were homogeneously mixed over space, or non-spatial model) was compared to a partially spatially explicit and discrete model (the spatial model), which was composed of differential equations and used geo-coded data (Euclidean distances between county centroids). While the spatial model accounted for intra-county and inter-county epidemic spread, the non-spatial model did not assess regional differences. A geo-coded dataset that resembled conditions favouring homogeneous mixing assumptions (based on the 2001 Uruguayan foot-and-mouth disease epidemic), was used for testing.

Significant differences between models were observed in the average transmission rate between farms, both before and after a control policy (animal movement ban) was imposed. They also differed in terms of daily number of infected farms: the non-spatial model revealed a single epidemic peak (at, approximately, 25 epidemic days); while the spatial model revealed two epidemic peaks (at, approximately, 12 and 28 days, respectively). While the spatial model fitted well with the observed cumulative number of infected farms, the non-spatial model did not ($P < 0.01$). In addition, the spatial model: a) indicated an early intra-county reproductive number R of ~ 87 (falling to < 1 within 25 days), and an inter-county $R < 1$; and b) predicted that, if animal movement restrictions had begun 3 days before/after the estimated initiation of such policy, cases would have decreased/increased by 23 or 26%, respectively.

Spatial factors (such as inter-farm distance and coverage of vaccination campaigns, absent in non-spatial models) may explain why partially explicit spatial models describe epidemic spread more accurately than non-spatial models even at early epidemic phases. Integration of geo-coded data into mathematical models is recommended.

Keywords: Foot-and-mouth disease; spatial mathematical model; reproductive number; Uruguay; movement restrictions.

1 Introduction

Foot-and-mouth disease (FMD) is a highly infectious illness caused by an aphthovirus that affects cloven-hoofed animals such as pigs, cattle, and sheep (Alexandersen et al., 2003; Kitching et al., 2005). The likelihood that FMD will start an epidemic outbreak depends on various factors that include the susceptibility of the livestock, the potential mode(s) of transmission, and the effectiveness of intervention efforts. Control efforts have been based, since 1911, on the concept of the basic reproductive number, introduced by Sir Ronald Ross (1911) and Kermack and McKendrick (1927). The basic reproductive number (or R_0) is defined as the number of secondary cases generated by a primary case when the virus is introduced in a population of fully susceptible individuals at a demographic steady state (Diekmann and Heesterbeek, 2000). That is, R_0 measures the power of a disease to invade a population under conditions that facilitate maximal growth. Once an outbreak starts, the number of susceptible livestock decreases either through loss of susceptibles (i.e., they get infected) or from the implementation of control measures such as slaughter or vaccination. When $R_0 > 1$, the epidemic progresses. When $R_0 < 1$, the epidemic dies out. The higher the R_0, the faster the infecting agent runs out of susceptible individuals (i.e., the faster it decreases).

However, the valid measurement of R_0 is problematic. To do so, models should assess the actual transmission mechanisms (causes that induce effects, that is, deterministic models), which would require data that: a) are likely to be unknown or biased (i.e., delayed case reporting and

under-reporting), and b) are likely to vary over space and even time (i.e., roads, farms, animal density, animal and human movement) (Rivas et al., 2004). Yet, simple deterministic models have been regarded to yield useful insights, generate intriguing hypotheses, and guide future research (Anderson and May, 1991). Elaborate deterministic models have been used to guide epidemic control policy (Ferguson et al., 2001).

One major assumption of deterministic models is that, in the early phase of an epidemic (especially when the disease is exotic and, therefore, all animals are susceptible; and when the replication cycle of the infecting agent is brief), the transmission is so rapid that, for practical purposes, the scenario where the epidemic develops may be regarded as "space-less": under those conditions, susceptible individuals may be regarded to be homogeneously mixed and in close contact (Keeling, 1999). The homogeneous mixing assumption characterizes non-spatial models (Koopman, 2004). That assumption leads to consider all infected cases as identical and, therefore, control policies based on the homogeneous mixing model tend to apply the same intervention in the same fashion (i.e., buffer zones of equal diameter, within which the same policy is applied, such as ring vaccinations) (Müller et al., 2000).

To assess the validity of homogeneously mixing-based, non-spatial models, at least two factors are needed: a) a spatially explicit model to be compared to, and b) geo-coded and temporal epidemic data. One approximation to provide a (partially) spatially explicit alternative model is to investigate R_0 while using spatial (local) data that consider the centroid-to-centroid distances among all county pairs where an epidemic takes place (Glavanakov et al., 2001). In addition, a geo-referenced and temporal dataset where an exotic infecting agent characterized by a short replication cycle infects a population lacking immunity (such as FMD affecting cat-

tle) is needed. At its earliest epidemic phase, that scenario would resemble a homogeneously mixed scenario (Rivas et al., 2003a).

Consequently, this study explored the validity of a non-spatial model in relation to a spatial model that estimated the local and regional disease transmission. For that purpose, a geo-referenced dataset based on the 2001 Uruguayan FMD epidemic was used.

2 Methods

2.1 Georeferenced and temporal epidemic data

Data from the FMD epidemic that took place in Uruguay in 2001 were obtained from public sources (MGAP, 2001; PAHO, 2001; European Commission DG [SANCO] reports # 3342/2001 and 3446/2001). The index case of this epidemic was reported on April 23, 2001 (epidemic day 1). Over 79 consecutive days, 1763 cases (infected farms) were reported (Figure 1). Details on this epidemic have been reported elsewhere (Rivas et al, 2003a, 2003b, 2004). A data-based simulation of the 2001 FMD epidemic in Uruguay is given in the supplementary materials.

Inter-centroid county distances among all Region's I county pairs ($n = 861$) were generated by Geographical Information Systems (GIS) software by retrieving first and linking later the polygon's centroid value of every county ($n = 42$).

2.2 Spatial epidemic model

The number of secondary outbreaks generated by a primary outbreak during its entire period of infectiousness was classified as internal (within counties) and external (across counties). Parameter values were estimated from data using least-squares fitting techniques. Parameter uncertainty was assessed using the stochastic temporal dependence of the cumulative number of outbreaks. Standard deviations for the estimated parameters were also calculated.

The epidemiological unit of analysis was the number of infected farms per county (Table 1). Farms were classified as susceptible (S), latent (L), infectious and undetected (I), and detected and isolated (J). A susceptible farm in county i (in contact with the virus) was regarded to enter the latent (uninfectious and asymptomatic) class (L) at the rate $\sum_{j=1}^{n} \beta_{ij} I_j$. In other words, the rate of infection was assumed to be proportional to the sum of the weighted prevalences of infected farms from all counties j. Hence, the transmission parameters β_{ij} measured the impact on county i from direct and indirect "contacts" between $i-$county and the $j-$county. These "contacts" may be the result of animal relocation or movement, from the sharing of milk routes (drivers as "mechanical" vectors or carriers), shared veterinarians or overlapping visitors (buyers, salesmen of farm products, etc. Sellers et al., 1971). Far away farms were assumed to be less likely to share the same veterinarians, milk trucks or visitors. It was assumed that the rate of transmission β_{ij} between farms in counties i and j decayed exponentially fast with the Euclidean distance of their respective county centroids. The elements of the "mixing" or "contact" matrix β_{ij} (Anderson and May, 1991) were therefore expressed as:

$$\beta_{ij} = \beta(t)\ e^{-qd_{i,j}}, \tag{1}$$

where $\beta(t)$ denotes the average transmission rate of infectious farms within each county at time t, d_{ij} is the distance between the centroids of counties i and j (Figure 1d), and the parameter q (1/km) which quantifies the extent of average local spread ($1/q$ can also be interpreted as the FMD mean transmission range). Small values of q lead to widespread influence, whereas large values of q support the hypothesis that local spread is the key. For simplicity, uniform mixing within each county was assumed, that is, $d_{ii} = 0$. It was also assumed that latently infected farms "progressed" towards the infectious class after a mean time of $1/k$ days and that infectious farms were detected and isolated from other farms at the per-capita rate α. That is, α is the average time required to detect and isolate an infected farm.

The above definitions and assumptions led to the following FMD model:

$$\begin{cases} \dot{S}_i &= -S_i \sum_{j=1}^{n} \beta_{ij} I_j \\ \dot{L}_i &= S_i \sum_{j=1}^{n} \beta_{ij} I_j - kL_i \\ \dot{I}_i &= kL_i - \alpha I_i \\ \dot{J}_i &= \alpha I_i. \end{cases} \tag{2}$$

The dot denotes time derivatives while S_i, L_i, I_i, and J_i denote the number of susceptible, latent, infectious, and isolated farms in county i ($i = 1, 2, ..., n$). The distribution of the number of farms per county is given in Table 1. The above system falls within the class of metapopulation models that have been used extensively to study ecological processes in heterogeneous patchy environments. In fact, the spatially dependent transmission rates $\{\beta_{ij}\}$ correspond to the

metapopulation patch connectivity index (Hanski, 1998) once we re-interpret d_{ij} as a measure of the influence of the landscape on migration (Moilane and Hanski, 1998). The elements of $\{d_{ij}\}$ here were set as "indices" that captured the effects of local transmission factors such as wind direction and animal heterogeneity within farms (dairy, beef, etc.). Here, the county connectivity d_{ij} was approximated by the distance between counties. The incorporation of a few time-dependent control/interventions measures led to the following modified model (see compartment diagram in Figure 2a) :

$$\begin{cases} \dot{S}_i &= -S_i(t)\sum_{j=1}^{n}\beta_{ij}(t)I_j(t) - \nu(t)S_i(t) \\ \dot{V}_i &= \nu(t)S_i(t) - V_i(t)\sum_{j=1}^{n}\beta_{ij}(t)I_j(t) - \mu(t)V_i(t) \\ \dot{L}_i &= (S_i(t)+V_i(t))\sum_{j=1}^{n}\beta_{ij}(t)I_j(t) - k(t)L_i(t) \\ \dot{I}_i &= k(t)L_i(t) - \alpha(t)I_i(t) \\ \dot{J}_i &= \alpha(t)I_i(t) \\ \dot{P}_i &= \mu(t)V_i(t) \end{cases} \tag{3}$$

where the classes S_i, L_i, I_i and J_i were defined as before. Susceptible farms in county i (S_i) are vaccinated at rate ν (V_i); vaccinated farms in V_i enter the protected class P_i at rate μ; vaccinated farms in county i that have not yet reached protective levels (class P) enter the latent (uninfectious and asymptomatic) class (L) at the rate $\sum_{j=1}^{n}\beta_{ij}I_j$. The total cumulative number of reported infected farms as a function of time is given by $C(t) = \sum_{i=1}^{n} J_i(t)$ while the daily number of new reported infected farms is given by $\dot{C(t)}$, that is by $\alpha(t)\sum_{i=1}^{n} I_i(t)$.

The dependence of parameters $\beta(t)$, $\alpha(t)$, $\nu(t)$, and $\mu(t)$ on time allow for the possibility of implementing control measures at different times (Chowell et al., 2004). For simplicity, these parameters were modelled as simple step functions

$$\beta(t) = \begin{cases} \beta_0 & t < \tau_m \\ \beta & t \geq \tau_m \end{cases} \tag{4}$$

$$\alpha(t) = \begin{cases} \alpha_0 & t < \tau_v \\ \alpha & t \geq \tau_v \end{cases} \tag{5}$$

$$\nu(t) = \begin{cases} 0 & t < \tau_v \\ \nu & t \geq \tau_v \end{cases} \tag{6}$$

$$\mu(t) = \begin{cases} 0 & t < \tau_v \\ \mu & t \geq \tau_v \end{cases} \tag{7}$$

where $\tau_m = 5$ is the epidemic day when movement restrictions were put in place and $\tau_v = 13$ is the time when mass vaccination started.

2.3 The reproductive number

Because there was not sufficient data to estimate the basic reproductive number (R_0), the *internal* reproductive number of county i, R_i^{int} was defined as the number of secondary outbreaks generated by an outbreak in county i within the same county after $t > 4$, $R_i^{int} = \beta N_i/\alpha$ where N_i denoted the number of farms in county i and $1/\alpha$ was the average time it took to identify infected farms. The *external* (across counties) reproductive number of county i, R_i^{ext}, was defined as the number of secondary outbreaks generated by an outbreak in county i in other counties, where $j = 1, 2, ..., n; j \neq i$. $R_i^{ext} = \sum_{j \neq i}^{n} \beta N_j \ e^{-qd_{ij}}/\alpha$, that is, it was given by the

additive contributions of the number of secondary cases (after the first intervention) in county i. Hence, the contributions were weighted by distance.

2.4 Spatially homogeneous model

In order to assess the role of spatial heterogeneity, a description of the corresponding spatially homogeneous version (null-model) follows. We set the homogeneous mixing assumption, $\beta_{ij} = \hat{\beta(t)}$ where

$$\hat{\beta(t)} = \begin{cases} \hat{\beta}_0 & t < \tau_m \\ \hat{\beta} & t \geq \tau_m \end{cases} \tag{8}$$

The corresponding system of nonlinear ordinary differential equations for the spatially homogeneous model becomes

$$\begin{cases} \dot{S}(t) &= -\hat{\beta(t)} S(t) I(t)/N - \hat{\nu} S \\ \dot{V}(t) &= \hat{\nu} S - \hat{\beta(t)} V(t) I(t)/N - \hat{\mu} V \\ \dot{L}(t) &= \hat{\beta(t)} (S(t) + V(t)) I(t)/N - \hat{k} L(t) \\ \dot{I}(t) &= \hat{k} L(t) - \hat{\alpha} I(t) \\ \dot{J}(t) &= \hat{\alpha} I(t) \\ \dot{P}(t) &= \hat{\mu} V(t) \end{cases} \tag{9}$$

where S, V, L, I, J, and P denote the total number of susceptible, vaccinated, latent, infectious, isolated, and protected farms, respectively. The parameters $\hat{\alpha(t)}$, $\hat{\nu(t)}$, and $\hat{\mu(t)}$ depend on time in the same manner as in the spatially explicit model.

2.5 Parameter estimation

The intrinsic growth rate quantified epidemic growth between successive time periods. The initial region-specific intrinsic growth rates r_i (i =1,2,3) were estimated under the assumption of exponential growth. That is, r (with units of 1/day) was estimated by assuming that the cumulative number of reported farms was proportional to $\exp(rt)$, where t is time (days). Solving for r, we obtained $r = (ln(y(t)) - ln(y_0))/t$, where ln denotes natural logarithm and y_0 is the number of outbreaks reported the during the first reporting day. The intrinsic growth rate in Region III was estimated using the cumulative number of outbreaks from May 02 to May 07, 2001. This window of time was chosen because of lack of cases prior to May 02 (Figure 2b).

The model parameters $\mathbf{\Theta}$ =($\beta(t)$, $k(t)$, $\alpha(t)$, $q(t)$, $\nu(t)$, $\mu(t)$) and the initial number of exposed and infectious farms ($E(0)$ and $I(0)$) were estimated from the cumulative number of reported farms (t_i, y_i), where t_i denotes the i^{th} reporting time (79 reporting days) and y_i is the cumulative number of reported farms by least-squares fitting to $C(t, \mathbf{\Theta})$ (the cumulative number of reported farms for our ordinary differential equation (or ODE) model with interventions (3) in Region I (where the epidemic started and the majority of outbreaks occurred). This gives a system of 5 equations per county (42 counties in Region I, or 210 differential equations). The farm density of each county is provided in Table 1. A language of technical computing (MATLAB, The MathWorks, Inc.) was used to carry out the least-squares fitting procedure. Initial conditions were chosen within the appropriate ranges ($0 < \beta < 100$, $1/5 < k < 1/3$, $1/12 < \alpha < 1/4$, $0 < q < 10$, $0 < \nu < 10$, $0 < \mu < 10$). Parameter optimization was carried out using the Levenberg-Marquardt method with line-search (More, 1977). This method was

implemented in MATLAB with the built-in routine `lsqcurvefit.m`. The cumulative number of reported farms $J(t)$ under a spatially homogeneous mixing ODE model (9) was fitted to data using also the same procedure described above.

The asymptotic variance-covariance $\mathbf{AV}(\hat{\mathbf{\Theta}})$ of the least-squares estimate for the spatially explicit model (3) was computed using a Brownian bridge error structure to model the stochastic temporal dependence of the cumulative number of outbreaks. The explicit formula used is

$$\mathbf{AV}(\hat{\mathbf{\Theta}}) = \sigma^{\mathbf{2}}\ \mathbf{B}(\mathbf{\Theta_0})\ \nabla_{\mathbf{\Theta}}\mathbf{C}(\mathbf{\Theta_0})^{\mathbf{T}}\ \mathbf{G}\ \nabla_{\mathbf{\Theta}}\mathbf{C}(\mathbf{\Theta_0})\ \mathbf{B}(\mathbf{\Theta_0}), \tag{10}$$

where $\mathbf{B}(\mathbf{\Theta_0}) = [\nabla_{\mathbf{\Theta}}\mathbf{C}(\mathbf{\Theta_0})^{\mathbf{T}}\ \nabla_{\mathbf{\Theta}}\mathbf{C}(\mathbf{\Theta_0})]^{-\mathbf{1}}$.

An estimate of $\mathbf{AV}(\hat{\mathbf{\Theta}})$ is

$$\hat{\sigma}^2\ \hat{\mathbf{B}}(\hat{\mathbf{\Theta}})\ \nabla_{\mathbf{\Theta}}\hat{\mathbf{C}}(\hat{\mathbf{\Theta}})^{\mathbf{T}}\ \mathbf{G}\ \nabla_{\mathbf{\Theta}}\hat{\mathbf{C}}(\hat{\mathbf{\Theta}})\ \hat{\mathbf{B}}(\hat{\mathbf{\Theta}}), \tag{11}$$

where $\hat{\mathbf{B}}(\hat{\mathbf{\Theta}}) = [\nabla_{\mathbf{\Theta}}\hat{\mathbf{C}}(\hat{\mathbf{\Theta}})^{\mathbf{T}}\ \nabla_{\mathbf{\Theta}}\hat{\mathbf{C}}(\hat{\mathbf{\Theta}})]^{-\mathbf{1}}$, $\hat{\sigma}^2 = \sum(y_i - C(t_i, \hat{\mathbf{\Theta}}))^2/(I_{1xn}\ \ \mathbf{G}\ \ I_{nx1})$ and $\nabla_{\Theta}\hat{\mathbf{C}}$ are numerical derivatives of $C(\hat{\mathbf{\Theta}})$. The error structure (Davidian and Giltinan, 1995) was also modelled by a Brownian bridge ($\mathbf{G}$). Here $\mathbf{G}$ is an n x n matrix with entries $G_{i,j} = (1/n)$ $\min(i,j) - (ij)/n^2$ where n is the total number of observations. $\mathbf{G}$ captures the higher variability in the cumulative number of outbreaks observed on the middle course of the epidemic as well as the smaller variability observed at the beginning and the end of the epidemic. Confidence intervals of 95% were computed using the asymptotic variance of our parameter estimates (diagonal elements of $\mathbf{AV}(\hat{\mathbf{\Theta}})$). The asymptotic variance-covariance $\mathbf{AV}(\hat{\mathbf{\Theta}_0})$ for the nonspatial (homogeneous mixing) model can be similarly computed using $J(t)$ in model (9) instead of $C(t)$.

The improvement in goodness of fit provided by the spatial model was compared to the nonspatial model by the stepwise F test (Jacquez, 1996). In fact, if $RSS_{spatial}$ denotes the residual sum of squares obtained from the spatial model, and $RSS_{nonspatial}$ is the corresponding sum of squares from the non-spatial (homogeneous mixing) model, then

$$RSS_{spatial} = \sum_{i=1}^{n=79} (y_i - C(t_i, \hat{\mathbf{\Theta}}))^2 \tag{12}$$

$$RSS_{nonspatial} = \sum_{i=1}^{n=79} (y_i - J(t_i, \hat{\mathbf{\Theta}_0}))^2 \tag{13}$$

The F test is the ratio of the decrease in the residual sum of squares, divided by the decrease in degrees of freedom ($p_{spatial} - p_{nonspatial}$), all divided by the mean residual sum of squares obtained from the spatial model ($RSS_{spatial}/(n - p_{spatial})$). That is,

$$\frac{(RSS_{nonspatial} - RSS_{spatial})/(p_{spatial} - p_{nonspatial})}{(RSS_{spatial}/(n - p_{spatial}))} \sim F_{(p_{spatial} - p_{nonspatial}),(n - p_{spatial})} \tag{14}$$

where $p_{spatial} - p_{nonspatial} = 1$ (the spatial model has only one additional parameter (parameter q) than the nonspatial model). When the above ratio is greater than the corresponding value of the F distribution for the significance level chosen then we would conclude that the spatial model significantly decreases the residual variance (Jacquez, 1996). Epidemic regions were identified in terms of the proportion of cases by use of the χ^2 test, conducted with statistical software.

3 Results

3.1 Regional epidemic growth rates

Three epidemic regions could be differentiated based on the percentage of all cases noticed in each region (Table 1 and Figure 2b, c). By epidemic day 79, Region I displayed 57% of all cases, while Regions II and III reported 32% and 11%, respectively ($P < 0.05$, χ^2 test). The initial intrinsic growth rates r shown in Regions I-III were 0.65, 0.35, and 0.19, respectively. After the 10th epidemic day, these growth rates decayed, becoming similar in all three regions (Figure 2b). Because Region I reported most cases throughout the epidemic and, consequently, it was the region most likely to display an environment that would correspond to the homogeneous mixing model, it was chosen for further analyses.

3.2 Model selection

When the non-spatial epidemic model was fitted to the cumulative number of Region I's infected farms, a systematic deviation was noticed during the first 20 epidemic days between fitted and observed epidemic data (Figure 3a). In contrast, when the cumulative number of Region I's infected farms was fitted using the spatial model, a close agreement was revealed (Figure 3b). The spatial model fit differed statistically significantly from that of the non-spatial model ($P < 0.01$, F-test). Additional differences between models were noticed when best-fit solutions were compared to the observed daily number of infected farms. While the non-spatial model indicated a single epidemic peak (taking place at, approximately, the 25th epidemic day),

the spatial model showed two epidemic peaks, occurring at epidemic days 10th and 28th , respectively (Figure 3c, d).

3.3 Comparison between models in terms of estimated parameters

The parameter estimates calculated by these models indicated both significant differences and similarities. The initial transmission rate was estimated by the non-spatial model to be $\hat{\beta}_0 = 0.77$ (S.D 0.04) in the first 4 epidemic days, while the spatial model estimated it as $\beta_0 = 0.33$ (S.D. 0.13). After the fifth epidemic day (start of movement restrictions), the transmission rate was estimated to be 0.49 (S.D. 0.08) and 0.10 (S.D. 0.03) by the non-spatial and spatial models, respectively (Tables 2 and 3). Other parameters did not differ beween models. Because the effects of control policy (vaccination) were regarded to occur at or after the epidemic peak, vaccination-related parameters were not analysed.

3.4 Applications for evaluation of control policy

Both models were concordant in indicating a significant decrease in the average transmission rate after a nation-wide animal movement ban was imposed on epidemic day 5. However, they differed markedly in the magnitude of that reduction. The non-spatial model indicated a reduction of 36%, whereas the spatial model estimated a reduction of 70% (Tables 2 and 3).

The spatial model also predicted that, if animal movement restrictions had been imposed 3 days before (or after) the estimated date of initiation of such policy (epidemic day 5), the total number of infected farms would have been reduced (or increased) by 23 or 26%, respectively (Figure 4). The spatial model also allowed us to estimate the intra- and inter-county

reproductive numbers, although it only included data reported since epidemic day 5 (after animal movement restictions had been imposed). The average internal reproductive number for Region I's individual counties was ~ 87, while the external (regional average) was ~ 0.82. We estimated that Region I's internal reproductive number decreased to less than 1 by epidemic day 25th .

4 Discussion

Because observational epidemiology is a discipline that does not facilitate the implementation of controlled experimental designs, model evaluation is constrained to use simulated scenarios. Data obtained from the 2001 Uruguayan FMD epidemic were retrospectively used to assess and compare the models here described. However, case reporting of actual epidemics is likely to include errors not limited to delayed reporting and under-reporting. Therefore, this study should not be construed as an assessment of the epidemic that took place in Uruguay in 2001, but as a model evaluation that used hypothetical (although realistic) geo-referenced and temporal epidemic data. That is, this study should be considered within the frame of the data here reported.

A contrast was noticed between early estimates of intra-county and inter-county R (87.20 and 0.82, respectively). At least the intra-county's R estimate was likely to reflect the influence of an assumption of the model, which was that all secondary (and later) cases derived only from those reported in the first replication cycle (primary cases), which is equal to say that no indirect transmission (i.e., through human movements or delivery routes, such as those of milk trucks) could coexist, when in fact such transmission mode was not prevented. Therefore, the

early estimates for intra-county transmission were probably over-estimated. However, within 25 epidemic days, the spatial model indicated a decrease of the intra-county R to < 1. It is suggested that the use of two R's may improve the analysis of epidemic dispersal by assessing simultaneously two scales: a) the micro or local scale (the intra-county R), and b) the regional scale (the inter-county R). If used together, these two scales might describe epidemic processes into four major types: a) a high R^{int} and low R^{ext}, b) a low R^{int} and high R^{ext}, c) a low R^{int} and low R^{ext}, and d) a high R^{int} and high R^{ext} types. A high R^{int} and low R^{ext}, type, as seen here, indicates that the force of infection fades with distance, which ultimately suggests that epidemic spread can only be sustained if the chances provided by long-distance connections coexist with favourable local conditions. Under such scenario, R^{int} must be very high for epidemics to progress (Holmes, 1997). Because the simultaneous use of two R estimates such as these has not been explored before, it was not possible to make comparisons to previously reported R values.

While the history of this scenario included the implementation of a national vaccination campaign (initiated on epidemic day 17th), which was assumed to take place over two weeks and require, at least, an additional week before antibody titers reached protective immunity (European Commission DG (SANCO) report # 3342/2001; Doel, 2003), that intervention was not a factor in the period prior to the epidemic peak (achieved before epidemic day 28, Fig 3c,d). Consequently, vaccination did not influence the only time frame within which conditions could resemble homogeneous mixing (when $R > 1$, Anderson and May, 1991; Brauer and Castillo-Chávez, 2001). Further potential sources of bias included the scale of the variables (infected farms aggregated at county level). More precise estimates could have been obtained

if geo-referenced data on all individual farms had been available (Rivas et al., 2004).

The parameter estimates generated by these models showed a good fit with previously reported data. For example, Hugh-Jones and Wright (1970) reported a latent period of 3-6 days, while 95% CI: 2.6-5.6 days was estimated here (Table 3). Keeling et al. (2001) and Ferguson et al. (2001) estimated the infectious period at 8 days, in agreement with our (95% CI) 6.3-8.3 days estimate (Table 3).

In the scenario under analysis the spatial model revealed a better fit than the non-spatial model. For example, a statistically significant difference was found between models in relation to their fit with observed (cumulative) number of infected farms (Fig 3a,b). In addition, while the non-spatial model only revealed a single epidemic peak, a double epidemic peak was indicated by the spatial model.

While the double peak indicated by the spatial model seemed to contradict the expectation for epidemic decline after epidemic day 10th (shown in Figure 2b) further supported by the rapid decrease of R^{int}, that finding could probably be explained by the vaccination implemented in Region I since or after epidemic day 17th and/or human movement (European Commission DG (SANCO) report # 3342/2001). The movement of vaccinators and vehicles across farms could have passively spread the virus among infected (but not clinical) cases and susceptible animals, resulting in a second, although brief, epidemic peak.

Because of the better fit displayed by the spatial model than the non-spatial model in relation to observed data, it is concluded that in the scenario under analysis (where conditions very closely resembled those based on homogeneous mixing), non-spatial models seem, nevertheless, inappropriate to accurately describe epidemic dispersal. Mathematically, this can be expressed

as differences due to non-random/non-uniform data distributions, which is equal to say that spatial autocorrelation (although not investigated in this study) most likely occurred in this dataset (Moran, 1950).

Likely reasons that may explain why non-spatial models are inappropriate to plan or monitor interventions (i. e., vaccinations) relate to factors such as the farm spatial network, intervention spatial coverage, and the percentage of animals and time required to synthesize specific antibody titers with protective levels (Figure 5). Two opposing forces determine the result of post-intervention outcomes. The outbreak is composed of factors related to the virus (including the virus incubation period and the virus infectivity period) and factors related to the spatial farm contact network. These factors can promote or prevent epidemic dispersal. The intervention can be viewed as a spectrum that ranges between two poles: a) vaccine efficacy, and b) vaccination impact. Vaccine efficacy is composed of vaccine homology, vaccine production safety and vaccine potency testing. In addition, the intervention may be influenced by: a) the spatial coverage (including the proportion of vaccinated farms or vaccination inter-herd coverage, and the proportion of vaccinated animals or intra-herd coverage), b) the proportion of animals exposed to susceptible (non-vaccinated) animals, c) the initial immune response (proportion of vaccinated animals that develop antibodies), d) animals age, and e) antibody titer decay. All these factors possess spatial expressions, which need to be accounted. The net efficacy of a vaccination campaign (vaccination impact) is a fraction of the original vaccine efficacy. Vaccination programs require not only to achieve certain global percentage of vaccinated animals (coverage) but also ensure that such level is evenly achieved, since pockets of unvaccinated animals within vaccinated farms may allow the virus to re-invade (Keeling, 1999). Yet, we

have failed to find literature reporting data on the spatial distribution of vaccination coverage. Similarly, spatial and temporal data on the percentage of animals reaching protective immunity is not available in the FMD-related literature. Antibody titers may decrease 7% or more after 4-6 weeks post-vaccination (Armstrong and Mathew, 2001; Woolhouse et al., 1996). Antibody titer decay is also influenced by the age of the host (Woolhouse et al., 1997). Therefore, both epidemic dispersal and intervention outcomes may be influenced by factors that are distributed over space in a non-random/non-uniform fashion. Absence of such data might explain the poor fit shown by the non-spatial model.

5 Conclusions

These analyses suggest that spatially explicit models are more likely to reflect local epidemic processes than non-spatial ones, even at early phases of epidemic dispersal. The integration of geo-referenced data at the lowest possible scale (i.e., farm-level data, as opposed to county-aggregated data) and mathematical models is recommended.

Acknowledgments

G. Chowell was supported by a Director's Postdoctoral Fellowship from Los Alamos National Laboratory and the Mathematical and Theoretical Biology Institute (MTBI). An earlier version of this manuscript is published as Los Alamos Unclassified Report LA-UR-04-5400.

References

Alexandersen, S., Zhang, Z., Donaldson, A.I., Garland, A.J., 2003. The pathogenesis and diagnosis of foot-and-mouth disease. J. Comp. Pathol. 129: 1-36.

Armstrong, R.M., Mathew, E.S., 2001. Predicting herd protection against foot-and-mouth disease by testing individual and bulk milk samples. J. Virol. Methods 97, 87-99.

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans, Oxford University Press, Oxford.

Brauer, F., Castillo-Chavez, C., 2000. Mathematical Models in Population Biology and Epidemiology, Springer-Verlag, New York.

Chowell, G., Hengartner, N.W., Castillo-Chavez, C., Fenimore, P.W., Hyman, J.M., 2004. The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. J. Theor. Biol. 229, 119-126.

Davidian, M., Giltinan, D.M., 1995. Nonlinear Models for Repeated Measurement data. Monographs on Statistics and Applied Probability 62. Chapman & Hall/CRC.

Diekmann, O., Heesterbeek, J.A.P., 2000. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, Wiley.

Doel, T.R., 2003. FMD vaccines. Virus Res. 91, 81-99.

European Commission (DG, SANCO, 3342, and 3456/2001) Health & Consumer Protection Directorate-General, Food and Veterinary Office on missions conducted in June and October, 2001. See:

<<http://europa.eu.int/comm/food/fs/inspections/vi/reports/uruguay/vi_rep_urug_3342-2001_en.pdf>>
<<http://europa.eu.int/comm/food/fs/inspections/vi/reports/uruguay/vi_rep_urug_3456-2001_en.pdf>>, accessed on June 06, 2004.

Ferguson, N.M., Donnelly, C.A., Anderson, R.M., 2001. The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions. Science 292, 1155-1160.

Glavanakov A., White, D.J., Caraco, T., Lapenis, A., Robinson, G.R., Szymanski, B.K., Maniatty, W.A. 2001. Lyme disease in New York State: spatial pattern at a regional scale. Am. J. Trop. Med. Hyg. 65, 538-545

Hanski, I., 1998. Metapopulation dynamics. Nature 396, 41-49.

Hugh-Jones, M.E., Wright, P.B., 1970. Studies on the 1967-8 foot-and-mouth disease epidemic. The relation of weather to the spread of disease. J. Hyg. (Lond). 68, 253-271.

Jacquez, J.A., 1996. Compartmental Analysis in Biology and Medicine. Third Edition. Thomson-Shore Inc., Michigan.

Keeling, M. J., 1999. The effects of local spatial structure on epidemiological invasions. Proc. Roy. Soc. B 266, 859-867.

Keeling, M.J., Woolhouse, M.E., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J., Grenfell, B.T., 2001. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science 294, 813-817.

Kermack, W. O., and McKendrick, A. G., 1927. A contribution to the mathematical theory of epidemics, Proc. Roy. Soc. A 115, 700-721.

Kitching, R.P., Hutber, A.M., Thrusfield M.V., 2005. A review of foot-and-mouth disease with special consideration for the clinical and epidemiological factors relevant to predictive modelling of the disease. Vet. J. 169, 197-209.

Koopman, J., 2004. Modeling infection transmission. Annu. Rev. Public Health 25, 303-326.

Ministry of Agriculture, Livestock and Fisheries (MGAP), Montevideo, Uruguay. Available at: `http://www.mgap.gub.uy`, accessed on August 01, 2001.

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17-23.

More, J. J., 1977. The Levenberg-Marquardt Algorithm: Implementation and Theory. Numerical Analysis, ed. G. A. Watson, Lecture Notes in Mathematics 630, Springer Verlag, pp. 105-116.

Moilane, A., Hanski, I., 1998. Metapopulation dynamics: effects of habitat quality and landscape structure. Ecology 79, 2503-2515.

Müller, J., Schnfisch, B., Kirkilionis, M., 2000. Ring vaccination. J. Math. Biol. 41,143-171.

Pan-American Foot-and-Mouth Disease Center (PAHO), Rio de Janeiro, Brasil. Available at `http://www.panaftosa.org.br/novo`, accessed on June 10, 2002.

Rivas, A.L., Tennenbaum, S.E., Aparicio, J.P., Hoogestein, A.L., Mohmed, H.O., Castillo-Chavez, C., Schwager, S.J., 2003. Critical response time (time available to implement

effective measures for epidemic control): model building and evaluation. Can. J. Vet. Res. 67(4), 307-311.

Rivas, A.L., Smith, S.D., Sullivan, P.J., Gardner, B., Aparicio, J.P., Hoogesteijn, A.L., Castillo-Chavez, C., 2003. Identification of geographic factors associated with early spread of foot-and-mouth disease. Am. J. Vet. Res. 64, 1519-1527.

Rivas, A.L., Schwager S.J. , Smith S., Magri A., 2004. Early and cost-effective identification of high risk/priority control areas in foot-and mouth disease epidemics. J. Vet. Med. B 51: 263-271.

Ross, R. 1911. The Prevention of Malaria (John Murray, London).

Sellers, R.F., Herniman, A.J., Donaldson, A.I., 1971. The effect of killing and removal of animals affected with foot-and-mouth disease on the amounts of airbone virus present in looseboxes. Br. Vet. J. 127, 358-364.

Supplementary materials. A movie of the 2001 foot-and-mouth epidemic in Uruguay. Available online at `http://math.lanl.gov/ gchowell/FMD/supplementary.htm`.

Woolhouse, M.E.J., Haydon, D., Pearson, A., Kitching, R., 1996. Failure of vaccination to prevent outbreaks of foot-and-mouth disease. Epidemiol. Infect. 116, 363-371.

Woolhouse, M.E.J., Haydon, D.T., Bundy, D.A.P., 1997. The design of veterinary vaccination programs. Vet. J. 153, 41-47.

World Organisation for Animal Health (OIE). Available at `<<http://www.oie.int>>`, accessed on June 10, 2002.

Table 1: Distribution of farm density and outbreaks of the 2001 foot-and-mouth disease epidemic in Uruguay over 79 epidemic days.

Region I					Region II					Region III				
State	Counties	N_j	Inf.	Tot.	State	Counties	N_j	Inf.	Tot.	State	Counties	N_j	Inf.	Tot.
Soriano	12	140	463	1682	Paysandu	13	121	64	1567	Artigas	12	118	34	1421
Colonia	18	151	362	2724	Salto	16	111	56	1783	Rivera	10	206	14	2064
Rio Negro	12	77	178	925	S. Jose	10	243	68	2430	C. Largo	16	196	26	2744
					Flores	9	91	62	816	Lavalleja	14	235	15	3296
					Florida	16	152	109	2436	Rocha	12	190	12	2284
					Tacuarembo	16	152	111	2427	T. y Tres	11	163	59	1797
					Durazno	15	136	92	2043	Maldonado	13	136	12	1773
										Canelones	23	141	25	3800

Counties, number of counties per state; N_j, mean number of farms per county; Inf., number of outbreaks per state; Tot., total number of farms per state.

Table 2: Parameter definitions and estimates obtained from least-squares fitting of nonspatial epidemic model (9) to the cumulative number of infected farms over time (days) in Region I (Figure 3a). All the parameters have units 1/ days.

Params.	Definition	Estim.	SD
$\hat{\beta}_0$	Average transmission rate between farms *before* mov. restrictions	0.77	0.04
$\hat{\beta}$	Average transmission rate between farms *after* mov. restrictions	0.49	0.08
$\hat{\alpha_0}$	Rate of detection of infected farms *before* mov. restrictions	0.16	0.07
$\hat{\alpha}$	Rate of detection of infected farms *after* mov. restrictions	0.14	0.02
$\hat{k}$	Rate of progression from latent to infectious state	0.26	0.07
$\hat{\nu}$	Vaccination rate of susceptible farms	0.16	0.04
$\hat{\mu}$	Rate at which vaccinated farms achieve protective levels	0.31	0.05

Table 3: Parameter definitions and estimates obtained from least-squares fitting of spatial epidemic model (3) to the cumulative number of infected farms (Figure 3b) in Region I. All the parameters have units 1/ days except for q whose units are 1/Km. * Small values of q lead to widespread influence, while large values support local spread. Great mobility and frequent interactions among farms would lead to small values of q.

Params.	Definition	Estim.	SD
β_0	Average transmission rate within counties *before* mov. restrictions	0.33	0.13
β	Average transmission rate within counties *after* mov. restrictions	0.10	0.03
α_0	Rate of detection of infected farms *before* mov. restrictions	0.14	0.02
α	Rate of detection of infected farms *after* mov. restrictions	0.14	0.02
k	Rate of progression from latent to infectious state	0.28	0.05
q^*	Positive constant quantifying the extent of local spread	1.03	0.10
ν	Vaccination rate of susceptible farms	0.25	0.09
μ	Rate at which vaccinated farms achieve protective levels	0.14	0.03

Figure captions

Figure 1:

(a) Daily and (b) cumulative number of farms reported as infected during the 2001 Foot and Mouth Disease epidemic in Uruguay. The epidemic reached its maximum of 66 outbreaks on day 33 (25 May 2001). By day 79 (10 July 2001) 1762 outbreaks had been reported. Data were obtained from public records of the Uruguayan Ministry of Livestock, Agriculture, and Fisheries (MGAP), the Pan-American Health Organization, and the World Organization for Animal Health (OIE). The periodic dips in the data tended to coincide with weekends. (c) Map of Uruguay with county divisions. (d) Distribution of intercounty (Euclidean) distances which were obtained using a Geographic Information System (GIS) software. The centroid of each county was used to compute Euclidean distances.

Figure 2:

(a) Schematic representation of the status progression for farms in a given county used to model the epidemic, as explained in the text. (b) The initial intrinsic growth rate r for Region I, II and III for the epidemic over 79 epidemic days. (c) Region I, II and III comprise 3, 7 and 8 Uruguayan states, respectively (see Table 1). The circle (Region I) denotes the site where the index case was reported.

Figure 3:

The cumulative and daily number of farms reported as infected in Region I (Figure 2c), where the epidemic started (23 April 2001) and most outbreaks occurred. Circles represent the ob-

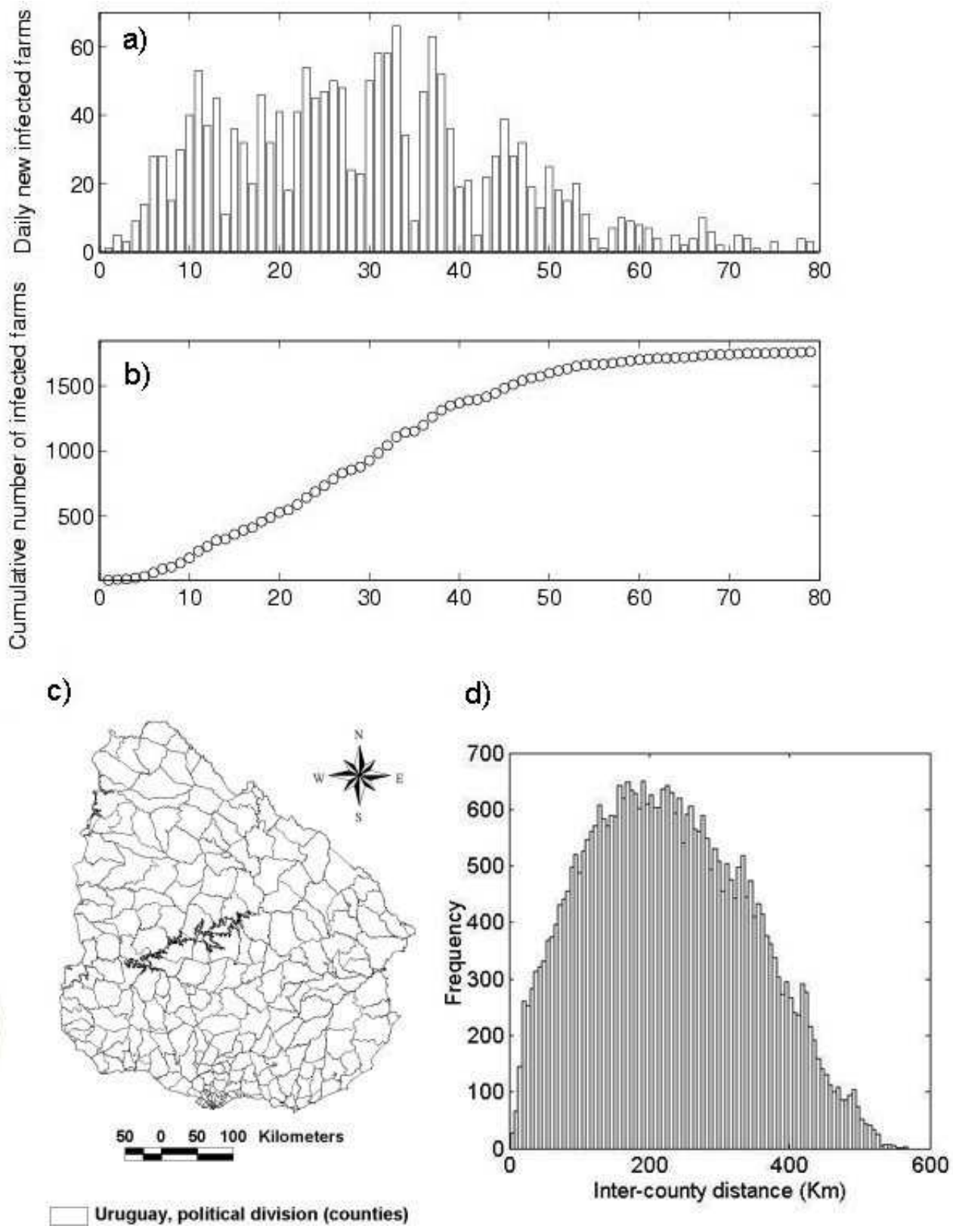
served data. The non-spatial model (9) fit is shown in (a) cumulative and (c) daily number of farms reported as infected in Region I. The spatial model (3) fit is shown in (b) cumulative and (d) daily number of reported outbreaks.

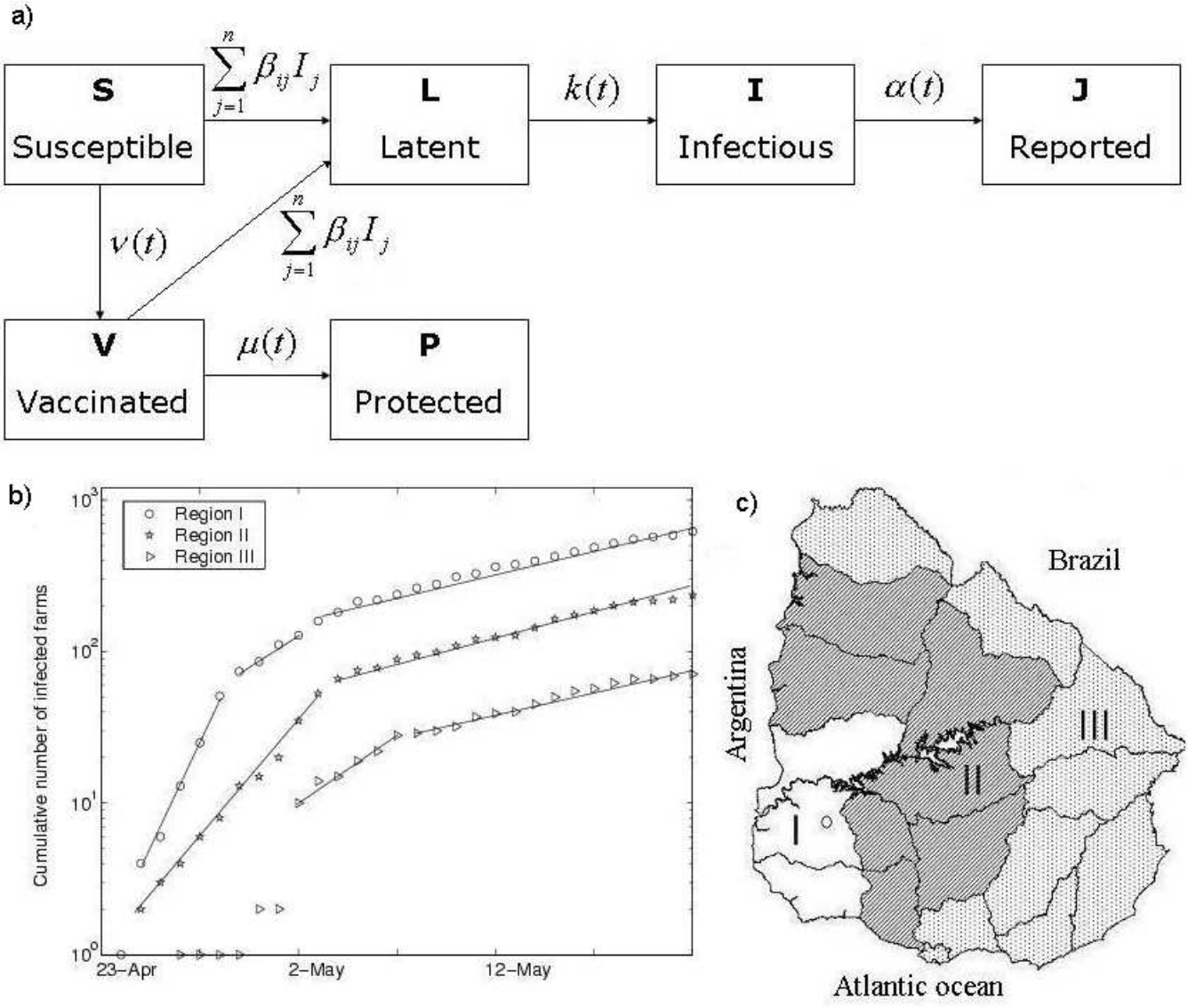
Figure 4:

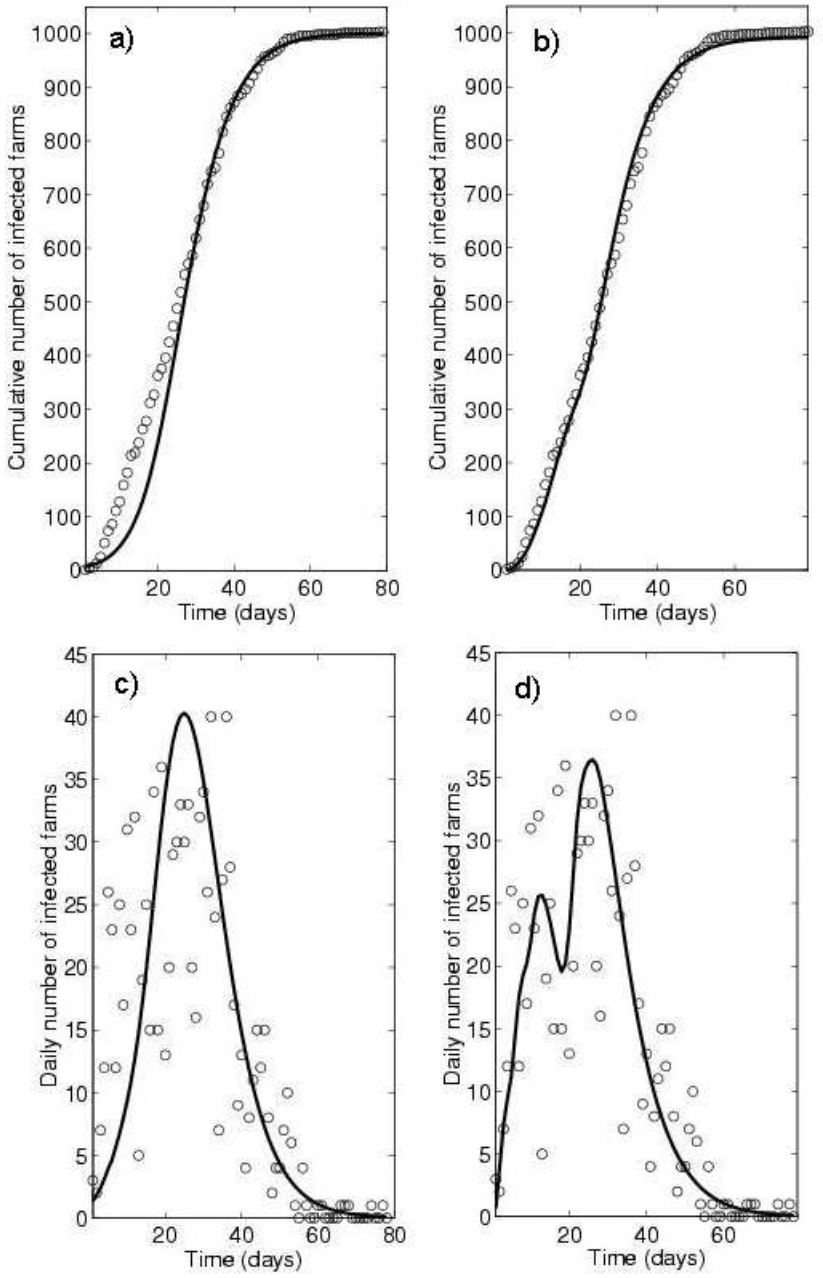
(a) The daily and (b) cumulative number of farms reported as infected in Region I (Figure 2c), where the outbreak started (23 April 2001) and most outbreaks occurred. Circles are the data, and the solid line is the best-fit solution of the spatial model equations (3) to the data by least-squares fitting (parameter estimates are given in Table 3). Two scenarios are shown: (dash-dash) movement restrictions with a 3-day delay and (dash-dot) 3 days before the actual date on which movement restrictions started.

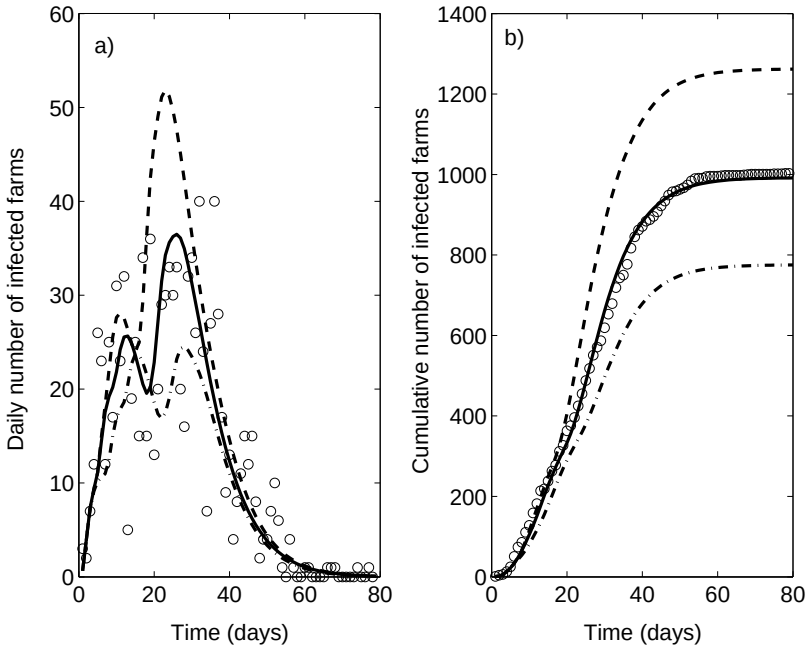
Figure 5:

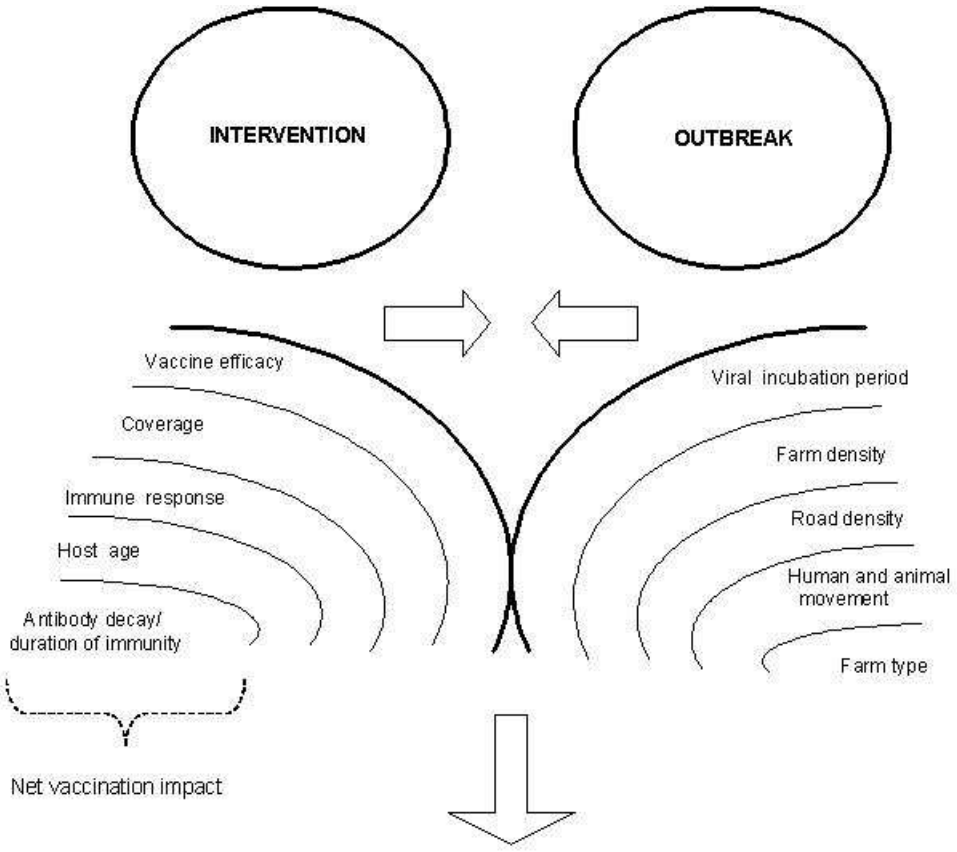
Concepts influencing outbreak-intervention (vaccination) interactions.


a)
Daily new infected farms
b)
Cumulative number of infected farms
c)
N
W
E
S
50 0 50 100 Kilometers
Uruguay, political division (counties)
d)
Frequency
Inter-county distance (Km)


a)
S
Susceptible
$\sum_{j=1}^{n} \beta_{ij} I_j$
L
Latent
$k(t)$
I
Infectious
$\alpha(t)$
J
Reported
$\nu(t)$
$\sum_{j=1}^{n} \beta_{ij} I_j$
V
Vaccinated
$\mu(t)$
P
Protected
b)
Region I
Region II
Region III
Cumulative number of infected farms
10^3
10^2
10^1
10^0
23–Apr
2–May
12–May
c)
Brazil
Argentina
Atlantic ocean
I
II
III


a)
b)
c)
d)
Cumulative number of infected farms
Daily number of infected farms
Time (days)
0
100
200
300
400
500
600
700
800
900
1000
5
10
15
20
25
30
35
40
45
20
40
60
80


a)
b)
Daily number of infected farms
Cumulative number of infected farms
Time (days)
Time (days)
0
10
20
30
40
50
60
200
400
600
800
1000
1200
1400
0
20
40
60
80



Outcome (epidemic dispersal vs. epidemic cessation)